Storytelling in the Age of Big Data:

Hong Kong Students' Readiness and Attitude towards Data Journalism

Fan Yang[1]
Ying Roselyn Du[2]

Abstract

This exploratory study identifies journalism students' attitudes, level of cognition and proficiency in data journalism (DJ). A survey of 121 college journalism students in Hong Kong, combined with in-depth interviews, found that: (a) while journalism students are eager to understand what is DJ and its practice, they do not have comprehensive knowledge of data collection, data analysis and interpretation; (b) computational tools are absent from current journalism curricula, which leads to students' misperception about data usage in news reporting; (c) while students have high willingness for learning DJ, about half of those surveyed expressed a dislike of data work. Gender, year in school and major course of study play a role in the students' varied perceptions of DJ. Male students mastered more DJ-related knowledge than their female counterparts. Those majoring in Chinese journalism show the least interest in DJ or learning its substance.

Keywords

Data journalism, journalism education, storytelling, data analysis, Hong Kong media, dislike of technology

[1] School of Journalism and Mass Communication, University of Wisconsin-Madison, Wisconsin, United States.
[2] Data & News Society at Hong Kong Baptist University, Kowloon Tong, Hong Kong.

Corresponding author:
Fan Yang, School of Journalism and Mass Communication, University of Wisconsin-Madison, United States.
E-mail: fanyang89@gmail.com

Introduction

Scholars have for decades discussed at length the close relationship between statistics as well as the application of mathematics in breaking news reports (Nguyen & Lugo-Ocando, 2016). Yet, the definitions of data journalism (DJ) as a concept and practice vary. Is DJ a new form of computer-assisted reporting (CAR), an extension of the social scientific element of 'precision journalism' (Folkerts, Hamilton, & Lemann, 2013; Meyer, 2002; Splendore et al., 2016) or a combination of both?

DJ arguably arises from Meyer's (2002) advocacy of precision journalism, which requires that journalists be well equipped with a working knowledge of social science research methods to uncover and interpret data and statistics underlying the story (Gunaratne, 1979; Weaver & McCombs, 1980). The advent of computing technology had made it easier for investigative journalists to sort, clean and interpret a range of data sets, hence the concept of CAR. DJ builds on the older practices of CAR and takes it further by engaging with data analysts, infographic designers and visualizers in producing and communicating the stories via the multimedia platforms.

Although DJ is regarded as the mantle of CAR, there are at least four different dimensions across these two overlapping fields: professionalism, openness, epistemology and visions of public (Coddington, 2015). While CAR was mostly driven by precision journalism (Meyer, 2002), DJ focuses more on audience interpretation, data visualization, large data size and response to the unbound information scale (Parasie & Dagiral, 2013). Initially, it was perceived that DJ primarily meant using data to report a story (Berret & Phillips, 2016). Others see DJ as a more advanced combination of mathematics, statistical computation (Nguyen, 2010), application of basic statistical analysis (Flew et al., 2012) and an array of techniques to find data.

Journalism that used big databases and statistics to break stories has since the 1980s won several Pulitzer prizes for general news or investigative reporting. For instance, *The Alabama Journal*'s usage of financial analysis to investigate the state's high infant mortality rate in 1988; and the series of investigative articles published by the *Wall Street Journal* on the admission preferences for white privileged students in elite American universities by Daniel Golden in 2004.

Today's environment of complex datasets and convergent media has made it an imperative for journalists to adapt to the era of information abundance driven by big data. This means normalizing the reporting of data-driven stories as part of the media work routine. For instance, open source data in leaked information by *Wikileaks* such as the Iraq War Logs, and the Panama Papers by the Panama-based law firm, Mossack Fonseca underscore the reality of big databases that require journalists be armed with a quick response, informed analysis and valid interpretation of complex datasets.

In Hong Kong, skills in data-based reporting are highly valued by middle-level journalists (Pacific Media Center, 2013). Irene Jay Liu, news editor for data at the

Thomson Reuters Hong Kong division, noted that DJ was a solution to the decline in newspaper circulation brought about by media convergence (ibid.). Yet, professional data journalists are prominently absent in the Hong Kong media industry. A cursory profile search of data journalist listed in Linkedin show that the top 25 DJs are mainly located in Western countries (Linkedin, 2016). Recent research on DJ have also focused mainly on the US experience (Plaue & Cook, 2015) even though Hong Kong boasts the most diversified, competitive and democratic media market (Freedom House, 2015) in the Asia-Pacific region.

Hong Kong has by far the most open society and enjoys the highest level of press freedom in Asia (Freedom House, 2015). University journalism programmes in Hong Kong are also highly ranked with Hong Kong Baptist University (HKBU) regarded as the top journalism school in Asia, according to *Asian Correspondent.*[1] However, research on the prevalence of DJ in the Hong Kong media industry is scarce, nor has any concrete study been done to examine Hong Kong students' existing knowledge structure of DJ, particularly in identifying the variance across gender, major and grades. This exploratory study aims to fill the knowledge gap on the teaching and training of DJ in Hong Kong. It surveyed 121 journalism majors in HKBU to first identify their internal knowledge structure of DJ—defined in this article as the students' conceptual/prior knowledge, interest and perception of DJ, and their career expectation—and second, to identify the areas for improvement in the teaching of DJ in university journalism programmes.

Data Journalism Education in Hong Kong

Traditional journalism education has predominantly focused on teaching professional skill sets and conventional news values. Journalism schools have a long history of overlooking the significance of teaching quantitative methods and investigating datasets in news reporting (Berret & Phillips, 2016; Russial & Santana, 2011; Yarnall et al., 2008). The experience of journalism schools in Hong Kong is no exception. With the growing significance of numbers and statistics in investigative work in the digital era, recent research have reiterated the need to integrate the teaching of traditional journalism with training in interpreting and analyzing complex datasets (Lorenz, 2011; Splendore et al., 2016).

To what extent has the neglect in teaching numeracy skills and data crunching affected or not affected journalism students' concept-knowing, interest, career expectation and prior knowledge in different specialized majors? To what extent would vocational journalism education that focused on specific skill sets affect the students' expectations of their future career in, for instance, DJ?

The fact is that journalism students are generally afraid of numbers (Maier, 2002; Wilby, 2007). Hence, it is important to see how their pre-existing knowledge structure could affect their continuing study of data-driven journalism. It has also been found that the participation of females and minorities in DJ is

smaller in university journalism programmes than in the industry as a whole (Berret & Phillips, 2016), which makes the role of gender in the DJ learning loop manifest. Previous research on DJ education in the United States noted that because there were more female students than male, female students tended to feel more vulnerable with data-related reporting assignments (Plaue & Cook, 2015).

Although professional journalists have set practical rules about how to train for reporting, analyzing and interpreting data (Gray, Bounegru, & Chambers, 2012), few studies have focused on journalism students' views on DJ. Our exploratory study of journalism students' perception of DJ education in Hong Kong is based on four blocks of perception: (a) 'concept-knowing', which means a general awareness of DJ; (b) 'interest' defined as students' personal attention to and liking for numbers and data-driven reports; (c) 'career expectation' refers to students' insight into the future of DJ practice; and (d) 'prior knowledge', which represents students' prior understanding and knowing about dealing with data-based reporting.

Our first two research questions are:

RQ1: What are the current situations of students' concept-knowing, interest, career expectation and prior knowledge of DJ in Hong Kong journalism schools?

RQ2: Do students' major, gender and year of study differentiate their concept-knowing, interest, career expectation and prior knowledge of DJ study?

Internal Knowledge Structure of Data Journalism

DJ education is currently identified as teaching a combination of traditional professionalism with a typical skills package (Tabary, Provost, & Trottier, 2016). Recent researches have provided guidelines to define the boundary between CAR and DJ but found that they all more or less bundle journalistic professionalism with data access (Lewis & Usher, 2013). The obvious difference between CAR and DJ is that the latter involves both investigative tools and a new way of story-telling in which visually and vividly presenting data is indispensable. So for DJ educators, the prominent issue is to incorporate professional values with techno-logy traits to train a qualified DJ practitioner in this field. A complete data report-ing course is usually thought to contain all stages of news production driven by data: collecting, analyzing, processing and visualizing data (Splendore et al., 2016). However, few studies have examined how students' perceive these four parts.

Based on the professional journalists' observation, the key to DJ education is to cultivate reporters' literacy in data collection, processing and analysis (Gray, Bounegru, & Chambers, 2012). One important feature of DJ is the rapid development of software applications, and how they could appropriately meet the needs of the complexity of reporting (Berret & Phillips, 2016). The importance of selecting techniques for teaching is based on a basic understanding of how much

prior knowledge our targeted students have. Some scholars think data visualization should be enhanced in the training while others emphasize the way programming works in geographic analysis (Appelgren & Nygren, 2014; Karlsen & Stavenlin, 2014; Parasie & Dagiral, 2013). However, the basic task of data-driven reporting is to access valuable raw data and then apply their logics and reasoning to the news that journalists need in order to help audiences to understand the stories. So, it is a requirement for a high caliber data journalist to not only master basic software skills but, more importantly, the ability to evaluate, analyze, question and visually present the data. To go from an abstract to a concrete vision, general understanding, tool usage and data processing are the three fundamental elements for a journalist, which is regarded as the internal structure of DJ study. Future DJ training should emphasize these three elements.

Because many free online tools applicable to data news reporting are now available, it would be wise to distinguish them by function. The difference is that each tool has a dominant advantage in the process of data application. Therefore, our third research question, in two parts, is:

RQ3: Is there any difference among students' perception about different types of DJ knowledge? If it is, in what aspect do they distinguish from each other?

Methods

Development of the Questionnaire

Before a large-scale online survey, two in-depth interviews were conducted to ascertain the key variables most relevant to current journalism students. Two undergraduate students from the School of Communication at HKBU were interviewed and their recorded conversations were transcribed as text. Questions were divided into three sections: 'General consulting', 'Specific knowledge' and 'Tool proficiency'. Keywords for each question were picked up as items for the questionnaire.

A total of 18 key dimensions were identified after a comprehensive review of these interviews. Four blocks of variables were developed based on the analysis. Cronbach's α was used to test their reliability. Guided by general principles of questionnaire design (Dillman, 2007), the 18 items, each measured on a five-point scale from 1 (strongly disagree) to 5 (strongly agree), were administered to students in the form of a three-page questionnaire.

The questionnaire consists of three parts. The first part measured students' concept-knowing, interest, expectation and prior experience. The second part was designed as a self-report combined with two open questions: 'What are the technical tools you know about for DJ that are not mentioned above?' and 'What do you have to say about DJ? The last section was three demographic variables to record subjects' study year, major and gender.

Measurement

Concept-knowing: Two items were tested on the scale. They are 'I know what DJ means' and 'I have never heard of DJ' (Cronbach's $\alpha = 0.75$).

   Interest: Three narratives were rated on the scale: 'I am interested in and pay attention to data-driven news stories'; 'I don't like dealing with numbers/data' and 'I don't like data-driven news stories' (Cronbach's $\alpha = 0.80$).

   Expectation: Two items were rated: 'Journalism students should learn how to produce data news' and 'Knowing DJ is an advantage for me in the future job market' (Cronbach's $\alpha = 0.72$).

   Prior Knowledge: 11 items were adopted to evaluate students' knowledge of DJ. Three categories were designed for all items. The first one is general experience: 'I have experience in DJ' and 'I know some popular tools used in DJ'. The second is knowledge about data usage: 'I know how to get data for news'; 'I know how to clean data using technical tools'; 'I know how to analyze data using technical tools' and 'I know how to visualize data using technical tools'. Finally, more detailed knowledge of tool use was examined: 'I know how to use Google Refine/Excel/ Google Fusion Tables/Tabula/to publish news stories online' (Cronbach's $\alpha = 0.86$).

Samples and Procedures

The online survey was conducted in September and October 2015. The sample was 121 college students from three undergraduate–postgraduate hybrid courses in journalism at HKBU. The HKBU was chosen because it has the largest journalism school with the most specialized fields in Hong Kong. Of the eight public universities in Hong Kong, HKBU has been widely recognized as the one with the most comprehensive array of professional journalism programmes[2] in Asia.

Data Analysis

A series of paired-samples *t*-tests and chi-square tests were conducted using SPSS 23.0 to examine the mean differences within and between groups. For RQ1, descriptive analysis was used to reflect a general outline of students' four cognitions towards DJ and for RQ3, paired-samples *t*-test and chi-square procedure were combined since it contains a series of subfield questions. For RQ2, since answers were recoded into three-level ordinal variables, chi-square procedures were taken for cross-tab analysis.

Results

Respondent Profile

Table 1 summarizes the students' demographic profiles. The majority of respondents are female Chinese journalism students from the second year.

Table 1. Profile of Respondents

(N = 121)

Demographic variables	Category	Percentage
Gender	Male	23.1
	Female	76.9
Years of study	First year	2.5
	Second year	57.1
	Third year	22.7
	Forth year	9.2
	Fifth year (master)	8.4
Major	Chinese journalism	41.5
	Financial journalism	33.9
	International journalism	24.6

Source: Authors' finding.

Current Situation of Data Journalism

For RQ1, items in the three blocks (concept-knowing, interest and career expecta-tion) were added and standardized into three options (no, yes and neutral) by computation. The label 'no' means not knowing the concept of DJ, having no interest in knowing and having no positive career expectation in this aspect. The label 'yes' represents the opposite argument to 'no' while 'neutral' speaks for 'has no attitude towards this issue'. The percentages for these three variables are presented in Figure 1.

The results show that more than 50 per cent of respondents at least knew or had heard of the concept 'DJ' and up to 84.3 per cent thought that DJ was advanta-geous to their future career and necessary for journalism study. However, less than half of the students had any interest in either numbers or data-driven news.

For RQ2, cross-tab and chi-square tests were used to see the variance of the four variables across gender, grade and major. The reason why two statistics were chosen is that for the variable of grade, not every item has more than 5 per cent of cases in the cell; so it is more accurate to see the differences by not converting their original scales.

As Table 2 shows, only gender plays a significant role in the difference of prior knowledge (χ^2 = 6.89, df = 1, $p < 0.05$). There were 28.3 per cent more male students with a high mastery of DJ-relevant knowledge than females. More details about different prior knowledge accumulation are illustrated in RQ3 and RQ4. However, gender differences had no significant impact on concept-knowing, interest and career expectation.

As regards major differences in the four variables, the statistics show signifi-cant differentiation in students' concept-knowing and interest: 65.3 per cent of students from a Chinese journalism major didn't like the concept of DJ; almost 80 per cent of students from the same major had no interest in either dealing with numbers and data or in data-driven news. Compared with Chinese journalism,
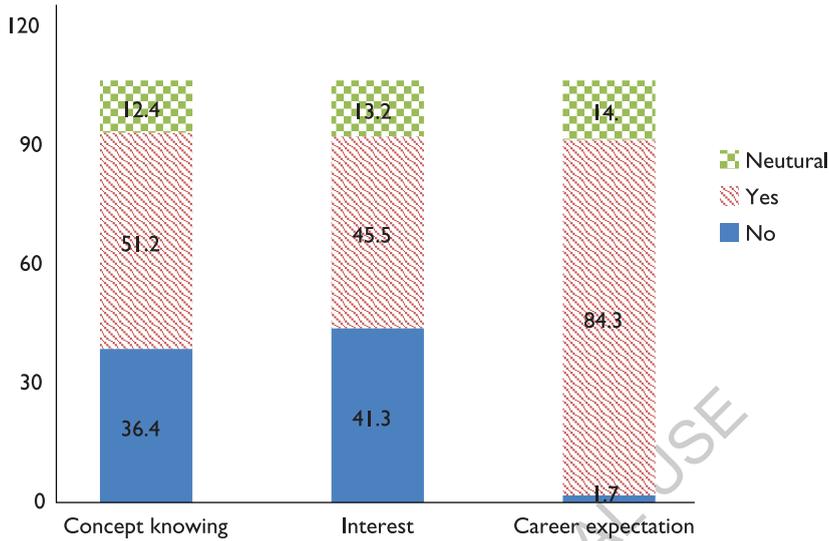
Figure 1. Percentages of concept-knowing, interest in and career expectation of data journalism

Source: Authors' finding.

Table 2. Gender differences in students' concept-knowing, interest in, career expectation of and prior knowledge of data journalism

Variables	Female n = 92		Male n = 28	
Concept-knowing	N	%	N	%
Not knowing	46	50	13	46.4
Knowing	46	50	15	53.6
$\chi^2 = .11$, df = 1, $p = .741$				
Interest				
Not interested	55	59.8	11	39.3
Have interest	37	40.2	17	60.7
$\chi^2 = 3.64$, df = 1, $p = .056$				
Career expectation				
No	16	17.4	3	10.7
Yes	76	82.6	25	89.3
$\chi^2 = 7.18$, df = 11, $p = .397$				
Prior knowledge				
Low ($\leq$ M = 26)	49	53.3	7	25
High (> M = 26)	43	46.7	21	75
$\chi^2 = 6.89$, df = 1, $p < .05$				

Source: Authors' finding.

respondents from financial journalism had the highest numbers in knowing DJ and processing numbers. Yet, the major line is that this did not impact students' career expectation and prior knowledge of data reporting (see Table 3).

Chi-square test was used to test the influence of grade (year of study) on the four variables. Table 4 shows that there exists a significant difference in concept interest (F = 8.90, $p < 0.001$) and career expectation ($F = 9.03$, $p < 0.001$). Data reveals that students from the third year had a slightly higher concept-knowing value than those in other grades (M_{year3} = 3.54) and students in the fifth year[3] had the highest expectations for DJ in their professional future (M_{year5} = 4.90).

For the observation of RQ3, we first added up the score of every item in three sections of knowledge. Then, we divided the three grades by the number of items to restore every knowledge section into the original five scales for comparison. A paired-samples *t*-test was processed to investigate whether there exists significant differences between every two sections and the result is presented in Table 5.

Results show that students only have significant differences either between general knowledge and tool usage knowledge or between general knowledge and data usage, whereas for knowledge between tool usage and data usage, their scores were not significantly different. That is to say, students' prior knowledge is different in a macro versus micro sense, but do not differ in mastery of concrete techniques.

For the knowledge of tool usage and data processing, Figure 2 shows the percentage of prior knowledge of tool usage among students. 'No' represents having

Table 3. Major differences in students' concept-knowing, interest in, career expectation of and prior knowledge of data journalism

Variables	Chinese journalism n = 49		Financial journalism n = 40		International journalism n = 29	
Concept knowing	N	%	N	%	N	%
Not knowing	32	65.3	11	27.5	16	55.2
Knowing	17	34.7	29	72.5	13	44.8
	χ^2 = 14.0, df = 3, $p < .05$					
Interest						
Not interested	39	79.6	10	25	17	58.6
Have interest	10	20.4	30	75	12	41.4
	χ^2 = 27.94, df = 3, $p < .001$					
Career expectation						
No	11	22.4	4	10	4	13.8
Yes	38	77.6	36	90	25	86.2
	χ^2 = 2.89, df = 3, $p = .409$					
Prior knowledge						
Low ($\leq$ M = 26)	22	44.9	17	42.5	16	55.2
High (> M = 26)	27	55.1	23	57.5	13	44.8
	χ^2 = 2.05, df = 3, $p = .562$					

Source: Authors' finding.

Table 4. Grade differences in students' concept-knowing, interest in, career expectation of and prior knowledge of data journalism

Grade	Concept-knowing	Interest	Career expectation	Prior knowledge (sum)
	M	M	M	M
Year 1 (n = 3)	3.50	4.17	4.50	32
Year 2 (n = 68)	2.90	2.69	3.78	26
Year 3 (n = 27)	3.54	3.39	4.19	28
Year 4 (n = 11)	3.41	3.75	4.18	26
Year 5 (n = 10)	3.50	4.05	4.90	25
	F = 2.29	F = 8.90	F = 9.03	F = .706
	p = .064	**p < .001**	**p < .001**	p = .589

Source: Authors' finding.

Table 5. Paired-samples t-test for three kinds of knowledge

Knowledge Comparison	Mean	Std. Error Mean	df	t
General vs. Tool usage	−.37	.81	114	−4.88**
General vs. Data usage	−.35	.76	111	−4.84**
Data usage vs. Tool usage	−.02	.64	107	.29

Source: Authors' finding.
Notes: (a) Items were standardized in a 5-point scale.
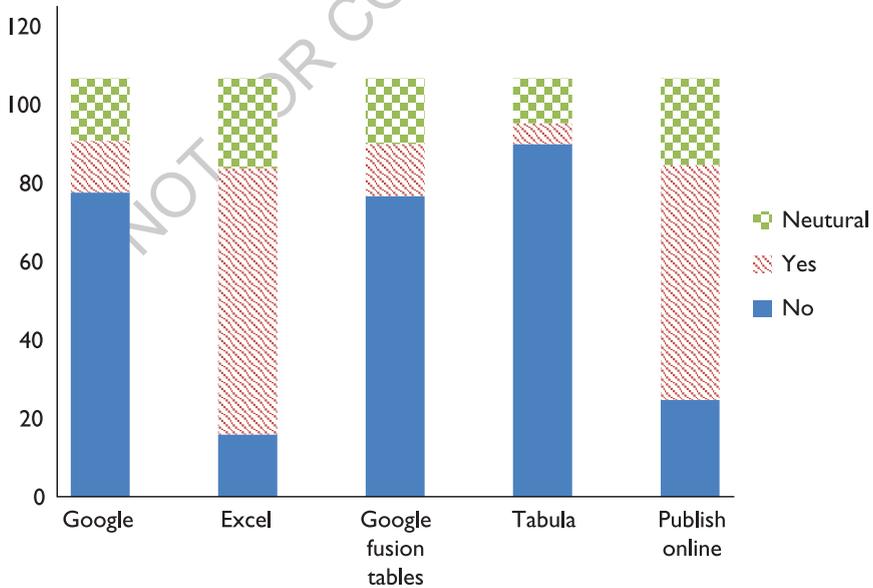   (b) **p < 0.01.

Figure 2. Percentage of prior knowledge for tool usage
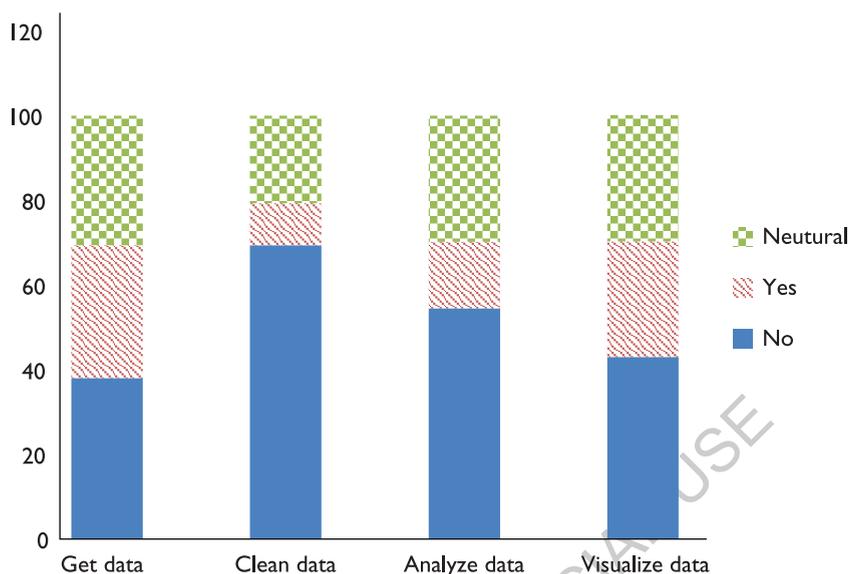
Source: Authors' finding.

Figure 3. Percentage of data processing in data journalism

Source: Authors' finding.

no prior knowledge in this software and 'Yes' means the opposite. It is obvious that most students knew how to use Excel (63.6 per cent yes) and publish news stories online (56.2 per cent yes), while they were less familiar with Tableau (84.3 per cent no), Google Refine (72.7 per cent no) and Google Fusion Tables (71.9 per cent no). For RQ3, Figure 3 shows that of data cleaning ranked as the top unknown skill (69.4 per cent) while getting data was the least unknown (38 per cent).

Discussion and Future Directions

Statistical analysis showed answers to the three research questions, and among these we can catch a glimpse of what current DJ education is like in the highly market-driven journalism field in Hong Kong. DJ is a widely discussed issue in media practice and in journalism schools. However, as the survey shows, journalism students do not truly show a deep understanding of DJ. More than half of the "students do not like numbers and data-driven news even though they acknowledge its value in job seeking. Second, 'gender, choice of major and prior' knowledge play an important part in the variance of interest in DJ. Male students appear to have more knowledge than their female counterparts, while Chinese journalism students dislike data-driven news the most.

We developed four indispensable elements to construct the structure of knowledge rooted in DJ studies. We found that the latest tools, such as Google Refine, Google Fusion Tables and Tabula, are essential to teaching DJ in the journalism programme. We also found that for most students, data cleaning using OpenRefine (formerly Google Refine) or Excel is the least known area of data processing.

The interviews and online survey presented in this study suggest that DJ education is like a double-edged sword. Although students recognize that knowing about DJ is a must for their career development and will be advantageous in their professional practice, thus, they need to acquire broader knowledge and practical training in DJ, they show minimal interest or possess minimal DJ skills in this area. The reason for this gap lies in insufficient training in DJ by the educators. Though there are adjunct lectures in journalism schools, DJ is somehow regarded by university administration as a 'cost' rather than an 'investment' (Nguyen & Lugo-Ocando, 2016). Consequently, it is hard for students to embrace comprehensive training in data analysis, not to mention its integration into journalistic values.

Although data usage has long been accepted in news reporting, criticism of its quality and authenticity continues (Shirky, 2014). Statistics are conceived with humans in mind, so it is the journalist's job to report the bad with the good, particularly in wake of the current deluge of open data. Journalistic bias prevails in modern trends as well as in old-fashioned practice; care should be taken so that journalists do not use statistics to reinforce their views and preconception of reality. Accountability, transparency and journalism ethics remain constant focus in DJ research (Bowles, Hamilton, & Levy, 2013; Cohen, 2011). Fact-checking, citing different sources and verifying information are traditional journalistic methods that need to be appropriately integrated with data processing and tool usage. An effective way must be found to design a mix of critical thinking, mastery of data and tool usage and programming concepts for a series of DJ curricula.

A future direction would be to combine in-class training with outside fieldwork. The differences between classroom direction and professional task specialization are worth investigating (Du, 2014). Considering the high speed of technological progress, the pace of updating college infrastructures, lecture skills and grading criteria should be observed and recorded carefully. For scholars, it is interesting to note the convergence and divergence of these two lines.

One limitation of this study is that we only have HKBU as the sample subject while there are two other universities in Hong Kong that offer journalism-related majors. The Journalism and Media Studies Center in Hong Kong University (HKU) offers a massive open online course (MOOC) in DJ while Chinese University of Hong Kong (CUHK) offers journalism at both undergraduate and postgraduate levels, although CUHK is known more for its academic orientation. Future surveys and interviews with journalism students in HKU and CUHK would well complement this study.

Another limitation is that in the survey of HKBU students did not touch on the data sourcing and programming, the two indispensable parts in data reports.

Future research could further examine the legal and ethical issues related to data acquisition and sourcing, both in the practical domain and in academia. The interplay of market forces, professional practice and class training should be investigated because what is instructed in journalism classes is better explicated with the latest trend in the field.

Interdisciplinary initiatives can also play a part in the training of data analysis and application. As DJ is a dynamic loop in practice, future research could combine the historical development of data reporting (from multimedia, boot camp to programming) with various feature stories to understand the effects of different kinds of storytelling. For scholars, the balance between data analysis, deadline pressures and editorial policy is also worth investigating.

Acknowledgements

Notes

1. https://asiancorrespondent.com/2011/04/top-10-journalism-schools-for-asian-students/ (Retrieved 5 September 2016).
2. The ranking of HKBU journalism education can be found from https://asian correspondent.com/2011/04/top-10-journalism-schools-for-asian-students/
3. We coded 'degree of master' into 'year 5', which means students classified as the fifth year students are pursuing a master's degree at HKBU.

References

Appelgren, E., & Nygren, G. (2014). Data journalism in Sweden: Introducing new methods and genres of journalism into 'old' organizations. *Digital Journalism, 2*(3), 394–405.

Berret, C., & Phillips, C. (2016). Teaching data and computational journalism. *Columbia Journalism School*. Retrieved April 2016, from https://www.gitbook.com/book/columbia journalism/teaching-data-computational-journalism/details

Bowles, N., Hamilton, J.T., & Levy, D. (2013). *Transparency in politics and the media: Accountability and open government*. London: I.B. Tauris.

Cohen, S. (2011). Shared values, clashing goals: Journalism and open government. *The ACM Magazine for Students, 18*(2), 19–22.

Dillman, D.A. (2007). *Mail and internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: Wiley.

Du, Y. R. (2014). What is needed versus what is taught: Students' perception of online journalism courses in Hong Kong. *Asia Pacific Media Educator, 24*(2), 225–237.

Flew, T., Daniel, A., Spurgeon, C.L., & Swift, A. (2012). The promise of computational journalism. *Journalism Practice, 6*(2), 157–171.

Folkerts, J., Hamilton, J.M., & Lemann, N. (2013). *Educating journalists: A new plea for the university tradition*. Retrieved 1 November 2016, from http://docplayer. net/16264121-Educating-journalists-a-new-plea-for-the-university-tradition.html

Freedom House. (2015). Freedom of press. Retrieved 1 November 2016, from https://freedomhouse.org/report/freedom-press/freedom-press-2016

Gray, J., Bounegru, L., & Chambers, L. (2012). The data journalism handbook. Retrieved 1 September 2016, from http://datajournalismhandbook.org

Gunaratne, S. A. (1979). Journalism teaching: Public affairs reporting, precision journalism and specialisms. *Australian Journalism Review, 2*(1), 4–10.

Karlsen, J., & Stavelin, E. (2014). Computational journalism in Norwegian newsrooms. *Journalism Practice, 8*(1), 34–48.

Lewis, S. C., & Usher, N. (2013). Open source and journalism: Toward new frameworks for imaging news innovation. *Media, Culture and Society, 35*(3), 602–619.

Linkedin. (2016). Data journalist. Retrieved 1 May 2016, from https://www.linkedin.com/title/data-journalist

Lorenz, M. (2011). Training data driven journalism: Mind the gaps Associationism. Retrieved 28 September 2014, from http://datadrivenjournalism.net/news_and_analysis/training_data_driven_journalism_mind_the_gaps

Maier, S.R. (2002). Numbers in the news: A mathematics audit of a daily newspaper. *Journalism Studies, 3*(4), 507–519.

Meyer, P. (2002). *Precision journalism*. Oxford: Rowman & Littlefield Publishers.

Nguyen, A. (2010). Harnessing the potential of online news: Suggestions from a study on the relationship between online news advantages and its post-adoption consequences. *Journalism, 11*(2), 223–241.

Nguyen, A., & Lugo-Ocando, J. (2016). The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism, 17*(1), 3–17.

Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: Computer-assisted reporters and programmer-journalists in Chicago. *New Media & Society, 15*(6), 853–871.

Pacific Media Center. (2013). NZ: Data journalism key for future journalists, says HK editor. Retrieved 19 June 2013, from http://www.pmc.aut.ac.nz/pacific-media-watch/nz-data-journalism-key-future-journalists-says-hk-editor-8333

Plaue, C., & Cook, L.R. (2015). *Data journalism: Lessons learned while designing an interdisciplinary course*. Paper presented at the 46th ACM Technical Symposium on Computer Science Education, February 2010, New York, US. Retrieved 1 September 2016, from http://dl.acm.org/citation.cfm?id=2677263&dl=ACM&coll=DL&CFID=863249976&CFTOKEN=78431109

Russial, J., & Santana, A. (2011). Specialization still favored in most newspaper jobs. *Newspaper Research Journal, 32*(3), 6–23.

Shirky, C. (2014). Last call: The end of the printed newspaper. *Medium*. Retrieved 1 March 2015, from https://medium.com/@cshirky/last-call-c682f6471c70

Splendore, S., Salvo, P.D., Eberwein, T., Groenhart, H., Kus, M., & Porlezza, C. (2016). Education strategies in data journalism: A comparative study of six European countries. *Journalism, 17*(1), 138–152.

Tabary, C., Provost, A. M., & Trottier, A. (2016). Data journalism's actors, practices and skills: A case study from Quebec. *Journalism, 17*(1), 66–84.

Utts, J. (2010, July). *Unintentional lies in the media: Don't blame journalists for what we don't teach*. Paper presented at the English International Conference on Teaching Statistics, Ljubljana, Slovenia.

Weaver, D.H., & Mccombs, M.E. (1980). Journalism and social science: A new relationship? *Public Opinion Quarterly, 44*(4), 477–494.

Wilby, P. (2007). Damn journalists and statistics *The Guardian*. Retrieved 12 February 2015, from http://www.theguardian.com/media/2007/nov/05/mondaymediasection.pressandpublishing

Yarnall, L., Johnson, J.T., Rinne, L., & Ranney, M.A. (2008). How post-secondary journalism educators teach advanced CAR data analysis skills in the digital age. *Journalism & Mass Communication Educator, 63*(2), 146–164.

Author's bio-sketch

Fan Yang is researching for her PhD at the School of Journalism and Mass Communication, University of Wisconsin-Madison, Madison. Her research focuses on hybrid media system, digital-era engagement and comparative politics. E-mail: fanyang89@gmail.com

Y. Roselyn Du, PhD, is founding principal coordinator of the multidisciplinary Data & News Society at HKBU. She was an awarding-winning journalist in China. She received her PhD at the University of North Carolina, Chapel Hill. Du has published in *Journalism and Mass Communication Quarterly* and *Journalism and Mass Communication Educator*, among others. E-mail: ydu@hkbu.edu.hk